International Conference on Intelligent Computing, Communication & Convergence

(ICCC-2015)

Conference Organized by Interscience Institute of Management and Technology,

Bhubaneswar, Odisha, India

Time Series Forecasting using Hybrid ARIMA and ANN Models based on DWT Decomposition

Ina Khandelwal*, Ratnadip Adhikari, Ghanshyam Verma

Department of Computer Science and Engineering, The LNM Institute of Information Technology, Jaipur-302031, India

Abstract

Recently Discrete Wavelet Transform (DWT) has led to a tremendous surge in many domains of science and engineering. In this study, we present the advantage of DWT to improve time series forecasting precision. This article suggests a novel technique of forecasting by segregating a time series dataset into linear and nonlinear components through DWT. At first, DWT is used to decompose the in-sample training dataset of the time series into linear (detailed) and non-linear (approximate) parts. Then, the Autoregressive Integrated Moving Average (ARIMA) and Artificial Neural Network (ANN) models are used to separately recognize and predict the reconstructed detailed and approximate components, respectively. In this manner, the proposed approach tactically utilizes the unique strengths of DWT, ARIMA, and ANN to improve the forecasting accuracy. Our hybrid method is tested on four real-world time series and its forecasting results are compared with those of ARIMA, ANN, and Zhang's hybrid models. Results clearly show that the proposed method achieves best forecasting accuracies for each series.

Keywords: Time serie forecasting; Discrete wavelet transform; ARIMA model; Artificial neural network; Zhang's hybrid model

* Corresponding author. Tel.:+91- 9414930740
  E-mail address: ina@lnmiit.ac.in

doi:10.1016/j.procs.2015.04.167

1. Introduction

Time series forecasting is a very active research topic in the domain of science and engineering. The primary objective of time series analysis is to develop a mathematical model that can forecast future observations on the basis of available data. Due to the difficulty in assessing the exact nature of a time series, it is often considerably challenging to generate appropriate forecasts. Over the years, various forecasting models have been developed in literature, of which the *Autoregressive Integrated Moving Average (ARIMA)* [1] and *Artificial Neural Network (ANN)* [1, 2] are widely popular. ARIMA models are well-known for their notable forecasting accuracy and flexibility in representing several different types of time series. But, a major limitation is their presumed linear form of the associated data that makes them inappropriate for complex nonlinear time series modeling [1]. ANNs successfully overcome this drawback of ARIMA models, but they have produced mixed results for purely linear time series. This means that neither ARIMA nor ANN is solely sufficient in modeling a real-world time series that almost always contains both linear as well as nonlinear correlation structures [1].

Wavelet transform has been applied in many engineering, signal processing, and statistical problems [3]. It is evident from the previous works that wavelet transform can considerably improve forecasting accuracies [4, 5]. Struzik [4] used wavelet decomposition in forecasting S&P index time series. Hsieh et al. [6] used *haar* wavelet decomposition for removal of noise from the stock price time series data with the aim to achieve more precise forecasts. Milidiu et al. [7] used haar wavelet together with a clustering algorithm to partition input data into different regions. Choi et al. [8] used a hybrid SARIMA and wavelet transform to forecast the sales time series. Conejo et al. [9] used wavelet transform to decompose electricity price time series and then applied ARIMA to forecast future prices. Wavelet transform has also been used to improve ANN forecasting accuracy [10, 11].

In this study, we propose a hybrid approach that adopts both ARIMA and ANN to adequately model the linear and nonlinear correlation structures of a time series, after a prior decomposition of the series through *Discrete Wavelet Transform (DWT)*. The proposed hybrid method is inspired from a similar concept, pioneered by Zhang [1] who pointed out that a real-world time series generally contains both linear and nonlinear patterns. Our approach accumulates the unique modeling strengths of ARIMA and ANN as well as the decomposition ability of DWT. The proposed method is implemented on four real-world time series and its forecasting accuracy is compared with that of ARIMA, ANN, and Zhang's hybrid approach in terms of two popular error measures.

The rest of the paper is outlined as follows. The next section describes the ARIMA, ANN, and Zhang's models. Section 3 presents our proposed hybrid approach. Section 4 reports the empirical results with conclusions in Sec. 5.

2. Time series forecasting models

Over the years, various time series forecasting models have been developed in literature. The random walk, autoregressive (AR), moving average (MA), and ARIMA are some widely recognized statistical forecasting models which predict future observations of a time series on the basis of some linear function of past values and *white noise* terms [1, 12]. As such, these models impose the inherent constraint of linearity on the data generating function. To overcome this, various nonlinear models have also been developed in literature. One widely popular among them is ANN that has many salient features [1, 13, 14]. Zhang [1] has rationally combined both ARIMA and ANN models in order to considerably increase the forecasting accuracy.

2.1. ARIMA model

The ARIMA models, pioneered by Box and Jenkins [12] are the most popular and effective statistical models for time series forecasting. These are based on the fundamental principle that the future values of a time series are generated from a linear function of the past observations and white noise terms. An ARIMA(p, d, q) model is mathematically expressed as follows:

$$\phi\left(B\right)\left(1-B\right)^{d} y_t = \theta\left(B\right)\varepsilon_t \tag{1}$$

Here, y_t, ε_t are the actual observation and white noise at time t, respectively; $\phi\left(B\right) = 1-\phi_1 B-\phi_2 B^2 -\cdots-\phi_p B^p$,

$\theta\;B\;=\;1+\theta_1 B+\theta_2 B^2+\cdots+\theta_q B^q$ are the lag polynomials, and B is the lag operator so that $By_t = y_{t-1}$. The constants p, q are the model orders, whereas $\phi_i,\;\theta_j\;\;i=1,2,\ldots,p;\;j=1,2,\ldots,q$ are the model parameters. The term d represents the degree of ordinary differencing, applied to make the series stationary. The appropriate orders of the ARIMA(p, d, q) model are usually determined through the Box-Jenkins model building methodology [1]. Due to the linearity restriction, an ARIMA may not be that effective for modeling a general real-world time series.

2.2. ANN model

ANNs constitute a very successful alternative to the ARIMA models for time series forecasting and have many distinguishing characteristics. One of them is the *universal approximation*, i.e. an ANN can estimate any nonlinear continuous function up to any desired degree of accuracy [13, 14]. A single hidden layer feedforward ANN with one output node is most commonly used in forecasting applications [1, 2]. A $p\times q\times 1$ANN has the following output:

$$y_t = \phi_0 + \sum_{j=1}^{q}\phi_j g\left(\theta_{0j} + \sum_{i=1}^{p}\theta_{ij} y_{t-i}\right) + \varepsilon_t \tag{2}$$

Here, $\phi_j\;(j=0,1,2,\ldots,q)$, $\theta_{ij}\;(i=0,1,2,\ldots,p;j=1,2,\ldots,q)$ are the weights, ϕ_0, θ_{0j} are the bias terms, and ε_t is the white noise. We take the *logistic function* [1, 13] as the hidden layer activation function g. So far ANNs have been very effective for modeling nonlinearly generated time series, but provided mixed results for linear problems [1].

2.3. Zhang's hybrid model

Evidently, neither ARIMA nor ANN is universally suitable for all types of time series. This is because almost all real-world time series contains both linear and nonlinear correlation structures among the observations. Zhang [1] has pointed out this important fact and has developed a hybrid approach that applies ARIMA and ANN separately for modeling linear and nonlinear components of a time series. According to Zhang, we have:

$$y_t = L_t + N_t \tag{3}$$

Where, y_t is the observation at time t and L_t, N_t denote linear and nonlinear components, respectively. At first, ARIMA is fitted to the linear component and the corresponding forecast $\hat{L}_t$ at time t is obtained. So, the residual at time t is given by $e_t = y_t - \hat{L}_t$. According to Zhang, the residuals dataset after fitting ARIMA contains only nonlinear component and so can be properly modeled through an ANN. Using p input nodes, the ANN for residuals has the following form: $e_t = f\;e_{t-1}, e_{t-2}, \ldots, e_{t-p}\;+\varepsilon_t$, where f is a nonlinear function, estimated by the ANN and ε_t is the white noise. If $\hat{N}_t$ is the forecast of this ANN, then the ultimate hybrid forecast at time t is obtained as:

$$\hat{y}_t = \hat{L}_t + \hat{N}_t \tag{4}$$

Through empirical analysis with three real-world time series, Zhang has found that his hybrid ARIMA-ANN method has achieved considerably better forecasting accuracies than both ARIMA and ANN models.

3. The proposed DWT based ARIMA-ANN hybrid method

Wavelet transform represents any function as a superimposition of a set of wavelets. As these functions are small waves, located in different times, the wavelet transform can provide crucial information about both time and frequency domains. Mathematically, a DWT can be represented as follows [5, 9, 11]:

$$\psi_{l,m}(t) = 2^{-l/2}\psi(2^l t - m) \tag{5}$$

Here, $l, m \in \square$ and $\int_{-\infty}^{\infty}\psi(t)dt = 0$. The terms l and m represent the scale factor and translation index, respectively.

In this study, we use DWT for obtaining a prior decomposition of a time series into high and low frequency components. The proposed approach consists of two phases, viz. *decomposition* and *reconstruction* [8, 15]. In the first phase, the series is decomposed into high (detailed) and low (approximate) pass filters, which respectively pick up the higher and lower frequency components of the series. In the next phase, the high and low frequency components are reconstructed through *inverse DWT (IDWT)* [5, 8]. After these two phases, an ARIMA is fitted to the reconstructed detailed part and forecasts are generated. Then, an ANN is fitted to the corresponding residuals together with the approximate part. Finally, the combined forecasts are obtained through adding these two component-wise forecasts. Our approach is elaborately presented in Alg. 1.

Algorithm 1. The proposed DWT based hybrid ARIMA-ANN forecasting algorithm

---

Assumption: $\mathbf{Y}$ (original series) = $\mathbf{Y}^{(LIN)}$ (linear part) + $\mathbf{Y}^{(NLIN)}$ (nonlinear part)

Inputs: The training dataset $\mathbf{Y}_{tr} = [y_1, y_2, \ldots, y_{N_{tr}}]^T$ and size N_{ts} of the testing dataset

Outputs: The combined forecast vector $\hat{\mathbf{Y}} = \left[\hat{y}_{N_{tr}+1}, \hat{y}_{N_{tr}+2}, \ldots, \hat{y}_{N_{tr}+N_{ts}} \right]^T$

Steps:

1. Apply DWT on $\mathbf{Y}_{tr}$ to decompose it as follows: [$\mathbf{A}$, $\mathbf{D}$]= DWT($\mathbf{Y}_{tr}$, 'filter')
    //$\mathbf{A}$ and $\mathbf{D}$ are the approximation and details parts, respectively

2. Obtain the reconstructions $\mathbf{Y}_{tr}^{ap}$ = IDWT $\left(\mathbf{A}, \text{'filter'} \right)$ and $\mathbf{Y}_{tr}^{ds}$ = IDWT $\left(\mathbf{D}, \text{'filter'} \right)$
    //It obtains the reconstructed approximation and details parts of $\mathbf{Y}_{tr}$

3. Find the appropriate ARIMA(p, d, q) model for $\mathbf{Y}_{tr}^{(ds)}$
4. Define $\alpha=(p+d+q)$, $N_{eff} = N_{tr} - \alpha$, and $\mathbf{Y}_{eff} = \mathbf{Y}_{tr}^{(ds)}(\alpha+1:N_{tr})$ *// $\mathbf{Y}_{eff}$ is the effective training dataset with size N_{eff}*
5. Initialize the residual vector $\mathbf{R}=\mathbf{0}(1:N_{eff})$ *// $\mathbf{0}$ is the zero vector*
6. Initialize $\mathbf{Y}_L= \mathbf{0}(1:N_{tr}+N_{ts})$, $\hat{\mathbf{Y}}_L= \mathbf{0}(1:(N_{eff}+N_{ts}))$, and $\hat{\mathbf{Y}}_{ts}^{(LIN)}= \mathbf{0}(1:N_{ts})$
7. Set $\mathbf{Y}_L(1:N_{tr})= \mathbf{Y}_{tr}^{(ds)}$
8. **for** k=1 to $(N_{eff}+N_{ts})$ **do**
9. $\hat{\mathbf{Y}}_L(k)$ = ARIMA $((p, d, q); \mathbf{Y}_L(k: \alpha+k-1))$ *//Obtain the kth ARIMA forecast through using $\mathbf{Y}_L(k: \alpha+k-1)$*
10.   **if** $k \leq N_{eff}$ **then**
11. $\mathbf{R}(k) = \mathbf{Y}_{eff}(k) - \hat{\mathbf{Y}}_L(k)$
12.   **end if**
13.   **if** $k > N_{tr}$ **then**
14. $\mathbf{Y}_L(k) = \hat{\mathbf{Y}}_L(k)$
15.   **end if**
16.   **end for**
17. $\hat{\mathbf{Y}}_{ts}^{(LIN)} = \hat{\mathbf{Y}}_L(N_{eff}+1:\text{end})$

{Estimation of the nonlinear part}

18. Initialize $\mathbf{Y}_N= \mathbf{0}(1:N_{eff}+N_{ts})$ and $\hat{\mathbf{Y}}_{ts}^{(NLIN)}= \mathbf{0}(1:N_{ts})$
19. Set $\mathbf{Y}_N(1:N_{eff}) = \mathbf{R} + (\mathbf{Y}_{tr}^{(ap)}(\alpha+1:\text{end}))$
20. Find the appropriate $i \times h \times 1$ ANN model for $\mathbf{Y}_N$
21. **for** k=1 to N_{ts} **do**
22. $\hat{\mathbf{Y}}_{ts}^{NLIN}(k)$ = ANN $\left((i,h,1); \mathbf{Y}_N\left(\overline{N_{eff}+k-i} : \overline{N_{eff}+k-1} \right) \right)$ *//Obtain the kth forecast through ANN*
23. $\mathbf{Y}_N(N_{eff}+k) = \hat{\mathbf{Y}}_{ts}^{NLIN}(k)$
24. **end for**
25. Obtain $\hat{\mathbf{Y}}=\hat{\mathbf{Y}}_{ts}^{(LIN)} + \hat{\mathbf{Y}}_{ts}^{(NLIN)}$

4. Results

We have conducted experiments with four real-world time series on MATLAB. The ARIMA and ANN models are implemented through the *Econometric* and *Neural Network* toolboxes [16], respectively. The model orders are estimated through several in-sample forecasting trials. The series are described in Table 1 and depicted in Fig. 1.

Table 1. Details of the time series datasets

| Time series	Description	Size (total, testing)	ARIMA	ANN
Lynx	Annual number of lynx trapped in Mackenzie river district (1821–1934) [17]	(114, 14)	(12, 0, 0)	7×5×1
Exchange rate	Weekly exchange rates from British pound to US dollar (1980–1993) [18]	(731, 52)	(0, 1, 0)	7×6×1
Indian mining	Monthly mining data of India (April, 1981–March, 1998) [19]	(204, 30)	(13, 0, 0)	12×9×1
US temperature	Average monthly temperature of Las Vegas, US (June, 1986–May, 2011) [17]	(300, 60)	(12, 0, 0)	12×6×1
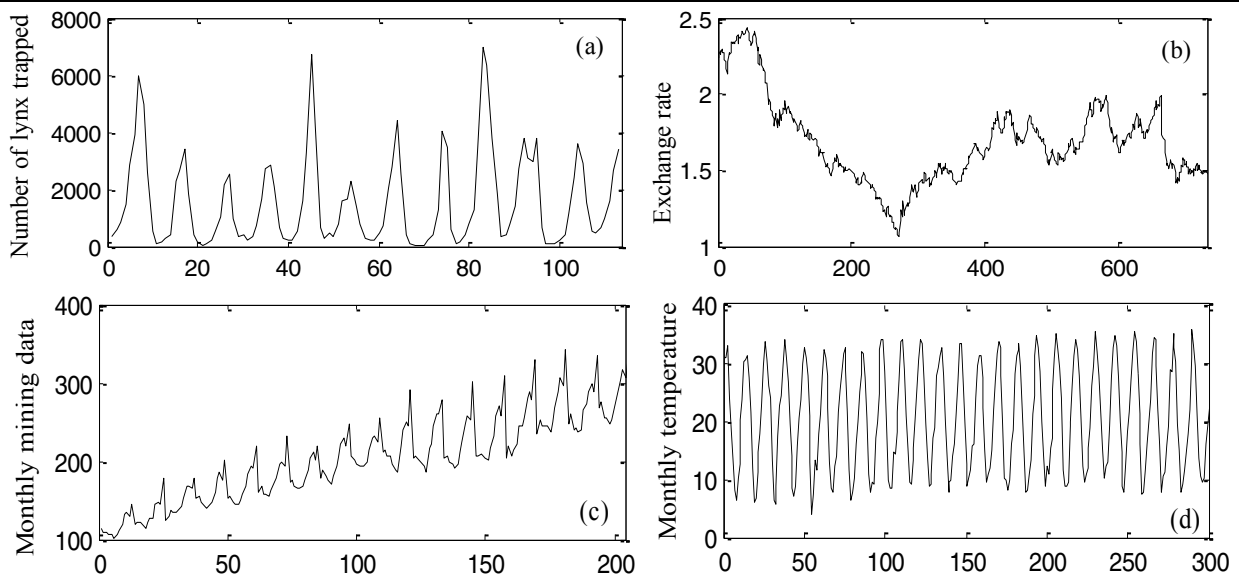
Fig. 1. The time plots of the series: (a) Lynx, (b) Exchange rate, (c) Indian mining, (d) US temperature

From the earlier works, it is evident that the *haar* and *daubechies* wavelets are most commonly used for forecasting applications [5, 6, 7]. To mitigate the problem of selecting the appropriate wavelet, in this study, we take the average of the forecasts through three wavelet transforms, viz. harr, db2, and db4. Table 2 presents the obtained forecasting results through ARIMA, ANN, Zhang's model, and the proposed hybrid method in terms of *Mean Squared Error (MSE)* and *Mean Absolute Percentage Error (MAPE)* [1, 14, 20]. From Table 2, it can be clearly seen that Zhang's model has achieved lower errors than both ARIMA and ANN and that our DWT based hybrid approach has outperformed all the three models through producing least errors. Figure 2 depicts the actual testing dataset (solid line) and its forecast (dotted line) through the proposed hybrid method for all four time series.

5. Conclusions

Achieving reasonably accurate forecasts of a time series is a very important yet challenging task. ARIMA and ANN are two widely popular and effective forecasting models. ARIMA assumes linear data generations function, whereas ANN is most suitable for nonlinearly generated time series. But, it is almost impossible to establish the exact nature of a series and a real-world time series most often contains both linear as well as nonlinear correlation structures. As such, in this paper, we have proposed a hybrid forecasting method that applies ARIMA and ANN separately to model linear and nonlinear components, respectively after a prior decomposition of the series into low

and high frequency signals through DWT. The final combined forecasts are obtained as the averages of the forecasts through harr, db2, and db4 wavelets. The empirical results with four real-world time series clearly demonstrate that the proposed method has yielded notably better forecasts than ARIMA, ANN, and Zhang's hybrid model.

Table 2. The obtained forecasting results

Methods	Lynx		Exchange rate[a]		Indian Mining		US temperature	
	MSE	MAPE	MSE	MAPE	MSE	MAPE	MSE	MAPE
ARIMA	0.01819	4.10977	18.4248	6.46017	68..7173	2.5035	4.66460	10.8206
ANN	0.01774	3.70555	10.8740	4.73684	76.3914	2.5897	17.9302	20.0074
Zhang's	0.01550	3.49327	10.6108	4.53310	68.2291	2.3644	3.83528	8.90731
Proposed	**0.00434**	**1.97038**	**9.07988**	**4.32839**	**43.8626**	**1.9411**	**1.64484**	**6.60378**

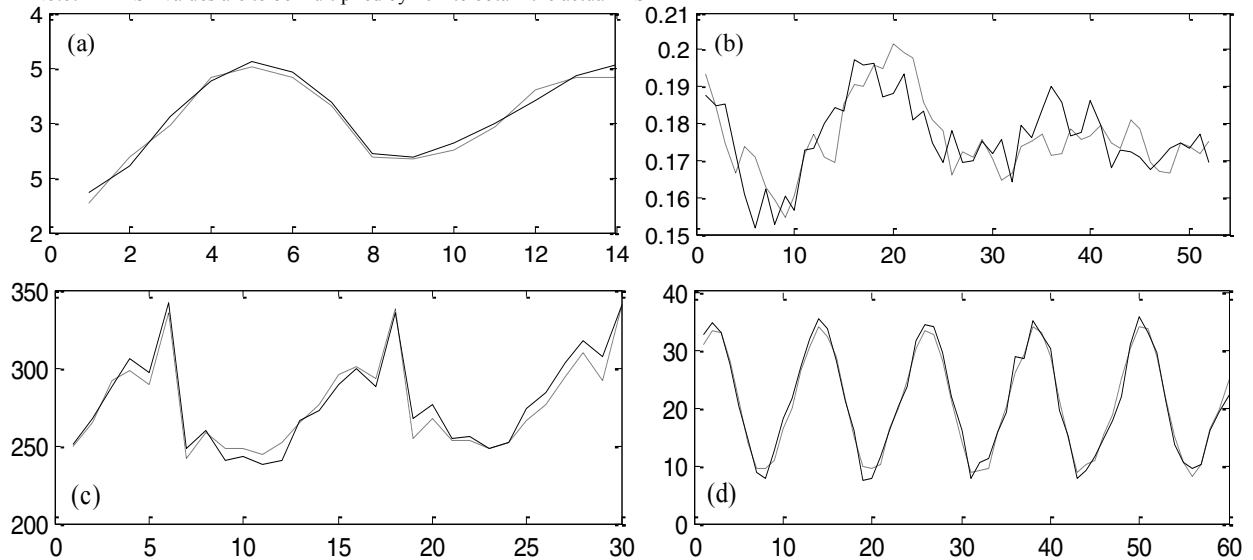[a] *Note*: All MSE values are to be multiplied by 10^{-5} to obtain the actual MSE

Fig. 2. Testing set and its forecast through the proposed method for: (a) Lynx, (b) Exchange rate, (c) Indian mining, (d) US temperature

References

1. Zhang, G.,P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing* 50, 159-175.
2. Zhang, G. P., Qi, M. (2005). Neural network forecasting for seasonal and trend time series. *European Journal of Operational Research* 160 (2), 501-514.
3. Chaovalit, P., Gangopadhyay, A., Karabatis, G., Chen, Z. (2011). Discrete wavelet transform based time series analysis and mining. *ACM Computing surveys*. http://doi.acm.org/10.1145/1883612.1883613. 43 (2), 0360-0300.
4. Struzik, Z. R. (2001): Wavelet methods in (financial) time-series processing. *Physica A: Statistical Mechanics and its Applications* 296 (1), 307-319.
5. Al Wadia, M. T. I. S., Tahir Ismail, M. (2011). Selecting wavelet transforms model in forecasting financial time series data based on ARIMA model. *Applied Mathematical Sciences* 5 (7), 315-326.
6. Hsieh, T. J., Hsiao, H. F., Yeh, W. C. (2011). Forecasting stock markets using wavelet transforms and recurrent neural networks: An integrated system based on artificial bee colony algorithm. *Applied soft computing* 11 (2), 2510-2525.
7. Milidiu, R. L., Machado, R. J., Renteria, R. P. (1999). Time-series forecasting through wavelets transformation and a mixture of expert models. *Neurocomputing* 28 (1), 145-156.
8. Choi, T. M., Yu, Y., Au, K. F. (2011). A hybrid SARIMA wavelet transform method for sales forecasting. *Decision Support Systems* 51 (1), 130-140.
9. Conejo, A. J., Plazas, M. A., Espinola, R., & Molina, A. B. (2005). Day-ahead electricity price forecasting using the wavelet transform and ARIMA models. *IEEE Transactions on Power Systems* 20 (2), 1035-1042.
10. Deka, P. C., Haque, L., Banhatti, A. G. (2012). Discrete wavelet-Ann approach in time series flow forecasting-a case study of Brahmaputra river. *International Journal of Earth Sciences and Engineering* 5 (4), 673-685.
11. Aussem, A., Murtagh, F. (1997). Combining neural network forecasts on wavelet-transformed time series. Connection Science 9 (1), 113-122.

12. Box, G. E. P, Jenkins, G. M. (1970). *Time series analysis, forecasting and control.* 3rd ed. Holden-Day, California.

13. Zhang, G., Patuwo, B. E., Michael, Y.Hu (1998). Forecasting with artificial neural networks: The state of art. *International Journal of Forecasting* 14 (1), 35-62.

14. Hamzaçebi, C. (2008). Improving artificial neural networks' performance in seasonal time series forecasting. *Information Sciences* 178 (23), 4550-4559.

15. Niu, C., Ji, L. (2012). A hybrid method based on wavelet analysis for short term load forecasting. *Journal of Convergence Information Technology* 7 (17).

16. Demuth, H., Beale, M., Hagan, M. (2010). *Neural network toolbox user's guide.* The MathWorks, Natic, MA, USA.

17. Data Market (2014). http://datamarket.com

18. Federal Reserve Bank of St. Louis (2014). http://wikiposit.org/uid?FRED.DEXUSUK

19. Open Government Data Platform, India (2014). http://data.gov.in

20. Adhikari, R., Agrawal, R. K. (2014). A combination of artificial neural network and random walk models for financial time series forecasting. *Neural Computing and Applications* 24 (6), 1441-1449.